

An Unsaturated Model for Request Mechanisms in WiMAX

Qiang Ni, *Senior Member, IEEE*, and Ling Hu

Abstract—Worldwide interoperability for microwave access (WiMAX) based wireless metropolitan area networks have recently received wide attention as they support high data rates with great quality of service (QoS) capabilities. In order to support QoS, bandwidth request (BW-REQ) mechanisms are suggested in the WiMAX standard for resource reservation. In this letter we propose a simple and accurate analytical model for the performance analysis of contention-based bandwidth request mechanisms. The accuracy of this model is verified with simulation results.

Index Terms—WiMAX, medium access control (MAC), bandwidth request, modeling.

I. INTRODUCTION

THE IEEE 802.16 WiMAX standard [1] specifies a new air interface for point-to-multipoint wireless broadband access, in which a base station (BS) serves a set of subscriber stations (SSs). Transmissions between BS and SSs are realized in a frame structure by means of time division multiple access (TDMA). To support duplexing, WiMAX defines both frequency division duplex (FDD) and time division duplex (TDD). Uplink transmissions from SSs to a BS are controlled by the BS through control messages, called uplink map (UL-MAP) messages transmitted in a prior downlink subframe. At medium access control (MAC) layer, the BS schedules the resources of uplink channel for initial ranging, bandwidth request (BW-REQ) and data transmissions. Various BW-REQ mechanisms are suggested in the WiMAX standard where contention-based BW-REQ is introduced for both non-real-time polling (nrtPS) and best effort (BE) services. With contention-based access, before each transmission attempt of a BW-REQ, an SS uniformly chooses a random integer number from the interval of $[0, W_j - 1]$, where W_j denotes its current contention window size [1]. The chosen value, also called a backoff counter, indicates the number of slots the station has to wait before the transmission of this request. For the first attempt, the window size starts with a minimal value W_0 . Upon each failure, the SS should double its window size, until reaching a maximum value. Above that, the SS can keep retransmitting the request with this maximum window size till reaching the retry limit and then dropping it. If a transmission is successful or stopped upon the retry limit, the window is then set back to W_0 .

Manuscript received October 9, 2009. The associate editor coordinating the review of this letter and approving it for publication was H. H. Chen.

The authors are with the Electronic and Computer Engineering Department, School of Engineering and Design, Brunel University, Uxbridge, Middlesex, UB8 3PH, UK (e-mail: Qiang.Ni@ieee.org).

This work was in part supported by the Brunel University's BRIEF award and the UK Engineering and Physical Sciences Research Council (EPSRC) under Grant No. EP/G070350/1.

Digital Object Identifier 10.1109/LCOMM.2010.01.091499

Furthermore, in WiMAX, contention scheme is allowed to work in a so-called grouping mode. In such case, contention is only performed for those SSs within a group to compete for BW-REQ transmission, where switching between groups is coordinated by the BS through polling. The grouping mechanism is suggested in WiMAX when available bandwidth is insufficient for a BS to individually poll many SSs.

Analytical modeling can help people to understand the performance of protocols under different scenarios and therefore to explore possible methods in improving the mechanisms. In the literature stochastic Markov chain models under saturated assumption have been proposed to analyze the performances of contention-based medium access in WiFi [2] and BW-REQ scheme in WiMAX [3]. However, networks do not typically operate in saturated conditions, particularly the generation of BW-REQs may be very infrequent. Hence an unsaturated model is needed for analyzing BW-REQ schemes. Furthermore, none of existing work considers grouping mode.

In the following we derive a simple analytical model that can compare various contention-based BW-REQ mechanisms (both grouping and no-grouping modes) in WiMAX with both saturated and unsaturated conditions.

II. AN UNSATURATED MODEL FOR CONTENTION-BASED BW-REQ IN WiMAX

Let us assume that there are one BS and n SSs in a WiMAX system, with each SS having one connection with the BS. For the simplicity of analysis, we assume that the total number of SSs, n , can be exactly divided into $g(\geq 1)$ groups with each group having n/g stations. Non-even group allocation, i.e. n is not dividable by g , will be our future work. Actually no-grouping (pure contention) scheme is a special case of grouping scheme, i.e. group number $g=1$. The BS polls each group in sequence. Only the stations within the group that is currently being polled, are allowed to decrement their backoff counters. For a given station in any group, we can model its backoff behavior with a discrete-time two-dimensional Markov chain denoted by $\{l(t), c(t)\}$ as depicted in Figure 1. The item $l(t)$ denotes the stochastic process of backoff stage at different contention window levels, and the item $c(t)$ represents the stochastic process of backoff timer at a given time slot t .

Note that WiMAX utilizes a TDMA based frame structure for BW-REQ and data transmissions. According to the WiMAX standard specification, after each transmission attempt the SS does not immediately start the backoff process for the next transmission attempt, instead it should wait till the beginning of the next frame. This behavior has not been modeled in [2][3]. In this work, we consider each WiMAX frame consists of a fixed number (denoted by K) of equal-sized slots for contention access. The duration of a slot is

chosen to be sufficient for one BW-REQ transmission. For example, if one SS completes its BW-REQ transmission at the k^{th} slot ($k = 0, 1, \dots, K-1$), it needs to wait for additional $(K-1-k)$ slots before it can start another backoff at the beginning of the next frame. As shown in Figure 1, we introduce in our Markov chain model the additional waiting states $\{-1, k\}$, $k = 0, 1, \dots, K-1$ to represent the above case when a station just completes a BW-REQ transmission, either successfully transmitted or discarded upon reaching the retry limit which is denoted by m . Since the time slot during which an SS can transmit is uniformly distributed at any of the K slots in each frame, the distribution of the number of additional waiting slots that an SS should wait is uniformly distributed over $[0, K-1]$.

In order to model both saturated and unsaturated conditions, the following behaviors are captured. On completion of a previous additional waiting period, if a BW-REQ is generated at the MAC queue (which occurs with probability p_r), a station can enter the first row of the backoff stage; otherwise (i.e. no request is generated) it has to wait for the next frame with the probability $1 - p_r$. When $p_r \rightarrow 1$, it models saturated conditions. In other cases, it models unsaturated conditions. Without loss of generality, we assume that BW-REQs arriving at the MAC of each SS follow Poisson distribution with arrival rate λ . Denote the average duration when a BW-REQ is generated till it leaves the MAC queue (either successfully transmitted or dropped) by $E(T)$, we can obtain the relation between p_r and the arrival rate λ as follows:

$$p_r = 1 - e^{(-\lambda E(T))} \quad (1)$$

Here, $E(T)$ also refers to the average delay of BW-REQs which will be determined later. Let p_c denote the probability that each BW-REQ collides. We assume this probability is independent of the number of the retransmissions which a request might have suffered. In the backoff stage, each time when the value of the backoff reaches zero the station makes a transmission attempt. If this transmission attempt is successful, the state will move to any of the $\{-1, k\}$ waiting states, and wait for the beginning of the next frame. On the other hand, when a collision occurs at the stage $j-1$, the backoff stage increases, and the new backoff counter is uniformly chosen in the range of $[0, W_j-1]$. At state $\{m, 0\}$ the BW-REQ will either be successfully transmitted or discarded by the station. After this state, the Markov chain goes back to the first row of the model (i.e. the additional waiting period) again.

Let $b_{j,k} = \lim_{t \rightarrow \infty} P\{l(t) = j, c(t) = k\}$ be the stationary distribution of the above Markov chain, with $j = -1, 0, 1, \dots, m$; $k = 0, 1, \dots, W_j - 1$, where $W_{-1} = K$. We denote the maximum contention window size by W_f , where f represents the maximum number that the window can be doubled. Therefore, we get $W_j = W_0 2^j$ when $0 \leq j \leq f$ and $W_j = W_f$ when $m \geq j > f$.

The normalization condition of this chain can be obtained:

$$\sum_{j=0}^m \sum_{k=0}^{W_j-1} b_{j,k} + \sum_{k=0}^{K-1} b_{-1,k} = 1 \quad (2)$$

Owing to the chain regularities, all probabilities in (2) can be expressed in $b_{0,0}$. Therefore, $b_{0,0}$ is obtained as follows:

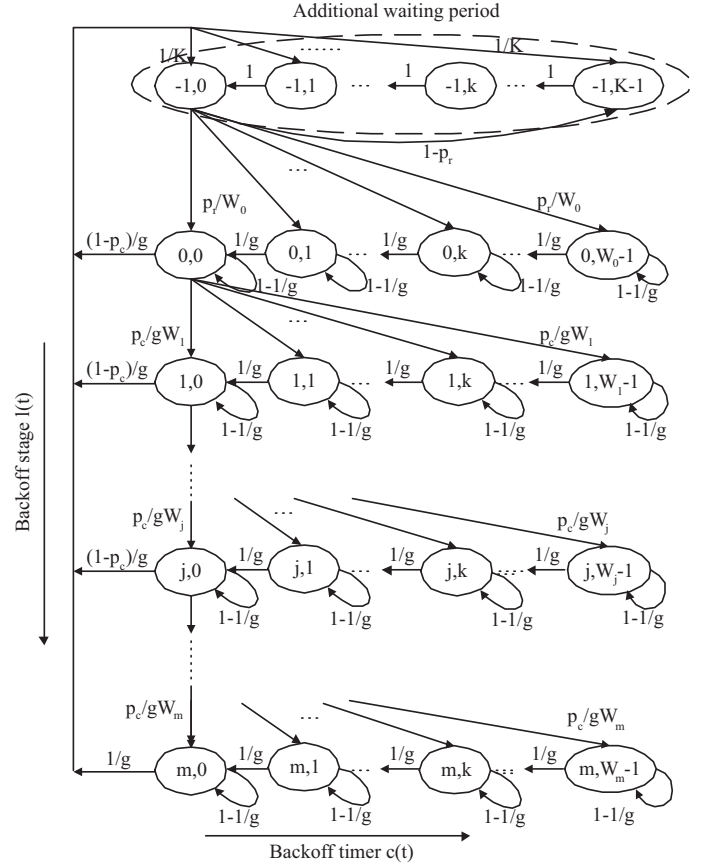


Fig. 1. An unsaturated model for no-grouping and grouping based BW-REQ access.

$$b_{0,0} = \frac{2gp_r(p_c - 1)}{gp_r[W_0 Z(p_c - 1) + p_c^{m+1} - 1] + (p_c - 1)[2K - p_r(K - 1)]} \quad (3)$$

where

$$Z = \frac{[(2p_c)^{f+1} - 1](p_c - 1) + 2^f(p_c^{m+1} - p_c^{f+1})(2p_c - 1)}{(2p_c - 1)(p_c - 1)} \quad (4)$$

By utilizing the Markov chain, the probability τ that an SS transmits a BW-REQ in a randomly chosen slot equals to:

$$\tau = \sum_{j=0}^m b_{j,0} = \sum_{j=0}^m p_c^j b_{0,0} \quad (5)$$

According to the Markov chain, the average delay of BW-REQ in slots can be calculated by:

$$E(T) = g \sum_{j=0}^m p_c^j \frac{W_j - 1}{2} + \frac{K}{2} \quad (6)$$

This considers the average waiting delay in K slots, the average backoff delay as well as the grouping effect.

Using (1) and (6), we obtain

$$p_r = 1 - e^{-\lambda \left[\frac{1}{2}g \left(W_0 Z - \frac{p_c^{m+1} - 1}{p_c - 1} \right) + \frac{K}{2} \right]} \quad (7)$$

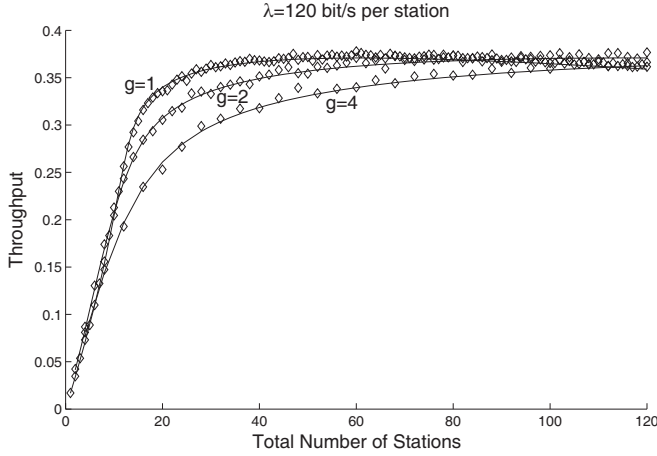


Fig. 2. Normalized throughput (grouping vs. no-grouping).

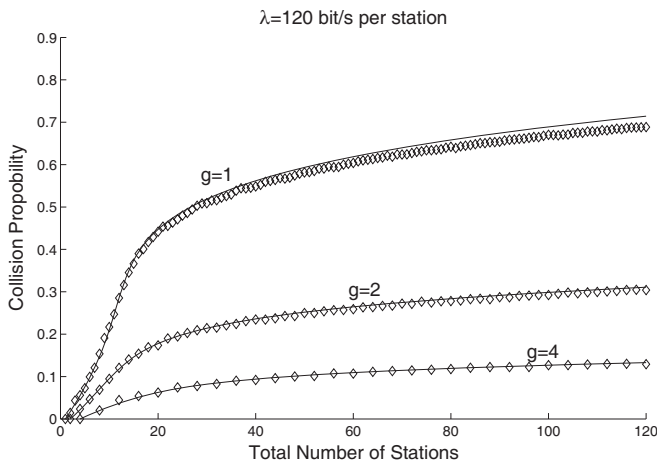


Fig. 3. Collision probability (grouping vs. no-grouping).

There are three unknown variables, τ , p_r and p_c . In order to solve such a non-linear system, we can use another condition: each station can transmit BW-REQ successfully with the probability $(1-p_c)$ only when all the remaining stations in the same group do not attempt to transmit during such a slot. Therefore it holds that:

$$1 - p_c = (1 - \tau)^{\frac{n}{g} - 1} \quad (8)$$

Hence we obtain three equations (5), (7) and (8) containing the aforementioned three unknown variables, which can be solved.

III. MODEL VALIDATION AND PERFORMANCE ANALYSIS

To verify the analytical results, we write Matlab codes to simulate the WiMAX systems. The simulation model implements the contention access protocols. In our following analysis, the frame length K is chosen to be 8 slots (i.e. $K=8$), the minimal window size W_0 is 16, and both f and m are set to 7.

We plot in Figure 2 and Figure 3 both the analytical model results (solid line) and the simulation results (diamond dot), including performance criteria of throughput and collision probability. Note that the simulation results are labeled with diamond points, with each point the mean value of 10 experiments¹. For every criterion, we vary the total number of stations with different types of group setting ($g=1,2,4$).

As shown in the both figures, the analytical model results match well the simulation results. The two figures show how normalized throughput and collision probability behave under different configurations. Figure 2 compares the throughput performance between grouping ($g=2,4$) and no-grouping ($g=1$). It demonstrates that when the system is unsaturated, no-grouping mode outperforms grouping mode. Furthermore, the more groups which we divide the stations into, the lower the normalized throughput we can achieve. The advantage of grouping rises when the system turns to be saturated.

Figure 3 demonstrates that with more groups, less stations will compete each other in a group and hence fewer collisions. Indeed, grouping reduces the risk of collisions. However, by dividing stations into groups, because of less number of stations per group, it also reduces the advantage of statistical multiplexing gain as compared to no-grouping mode. Therefore, there is a tradeoff in using grouping mode. When system load turns high, the risk of collisions gets high. The gain by reducing collisions is higher than the loss on statistical multiplexing gain, and hence grouping is recommended. On the other hand, when system is lightly loaded, grouping is not recommended.

IV. CONCLUSION

In this paper, we presented a simple analytical model which can predict the performance of contention-based BW-REQ mechanisms in WiMAX. Due to page limit, we presented only the estimation of throughput and collision probability of BW-REQs under saturated and unsaturated conditions. The model is verified by simulation results. The results show that grouping mechanism is helpful considering throughput when system load is high, but it is not applicable when system is lightly loaded. The proposed model can be used to design efficient operation of BW-REQ mechanisms under various network conditions.

REFERENCES

- [1] IEEE Std 802.16-2005, "IEEE Standard for Local and Metropolitan Area Networks," Part16: Air Interface for Fixed and Mobile Broadband Wireless Access Systems," Dec. 2005.
- [2] G. Bianchi, "Performance analysis of the IEEE 802.11 distributed co-ordination function," *IEEE J. Sel. Areas Commun.*, vol. 18, no. 3, pp. 535-547, Mar. 2000.
- [3] J. He, K. Guild, K. Yang, and H. H. Chen, "Modeling contention based bandwidth request scheme for IEEE 802.16 networks," *IEEE Commun. Lett.*, vol. 11, no. 8, pp. 698-700, Aug. 2007.

¹In each experiment, the time duration of 1000 periodic frames is chosen.